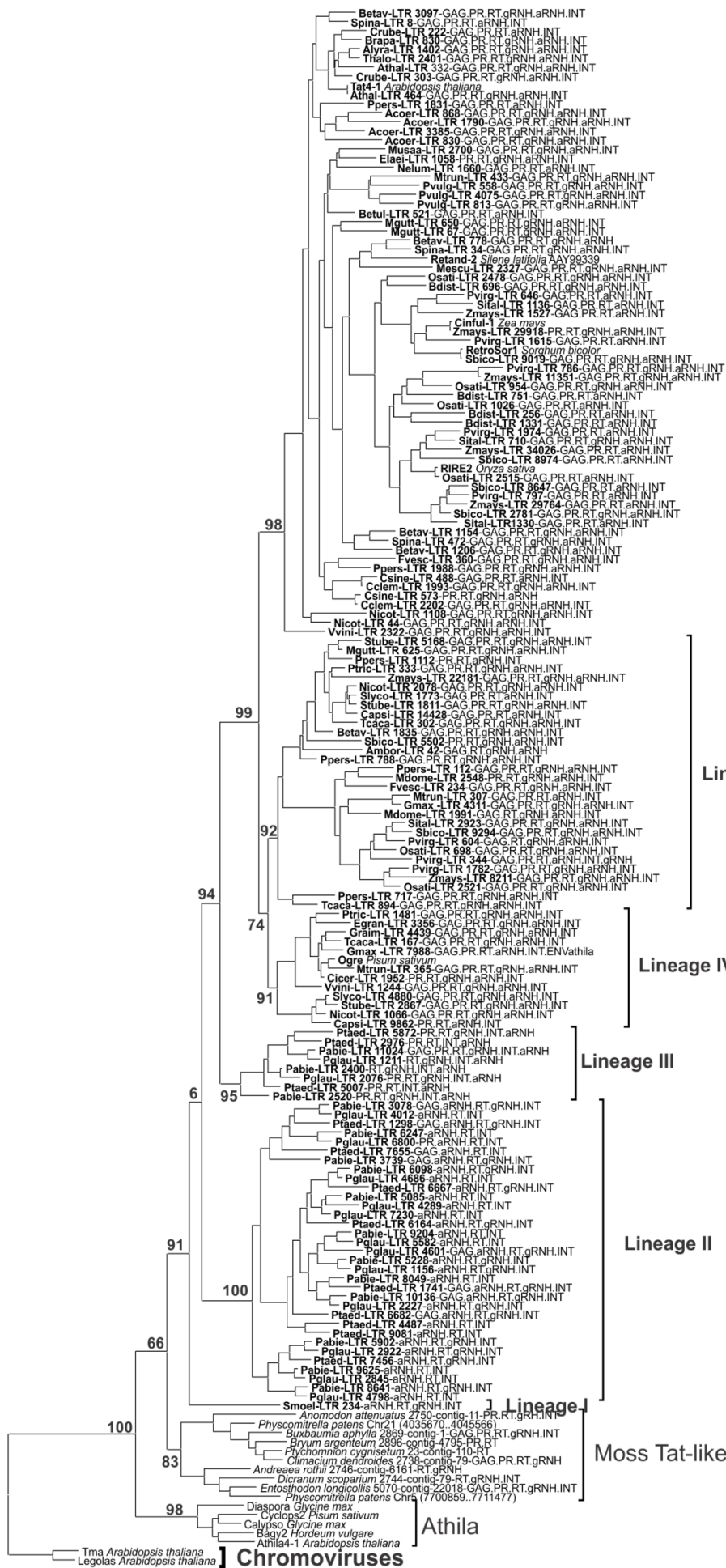


**Figure S1.** Maximum-likelihood tree based on the amino acid sequences of RT of diverse LTR retrotransposons, retroviruses, caulimoviruses and DIRS elements. The different types of retroelements are denoted with brackets to the right of the tree. Statistical support was evaluated by the approximate likelihood-ratio test (aLRT) and is shown at corresponding nodes of the tree. The names of all aRNH-containing LTR retrotransposons are in bold. The names of aRNH-containing elements identified in the study consist of the host genome acronym and the index number of the element, for example 'Smoel-LTR 234'. For species acronyms see Table S1. The names of moss Tat-like elements contain the name of the genome and the reference to the genomic sequence where the element was identified, for example '*Physcomitrella patens* Chr21 (4035670..4045566)'. The elements that were retrieved from GyDB were designated using its GyDB name and the name of host genome, for example, 'Ogre *Pisum sativum*'. The names of the elements that were retrieved from Repbase or Retrobase contain reference to corresponding database and the name of the host genome, for example, 'DIRS-1 BF *Branchiostoma floridae* Repbase'. Finally, the RT sequences that were retrieved from GenBank are designated using the genome name and corresponding accession number, for example, '*Chrysemys picta bellii* XP 005308735'.



Lineage VI

Athila/Tat

Lineage V

Lineage IV

Lineage III

Lineage II

Lineage I

Moss Tat-like

Athila

Chromoviruses

**Figure S2.** Maximum-likelihood tree based on the amino acid sequences of RT from representatives of Tat LTR retrotransposons, Athila elements and chromoviruses. Six lineages of aRNH-containing Tat LTR retrotransposons are denoted by latin numbers I-VI and the names of the elements are in bold. Moss Tat-like elements, which are phylogenetically close to aRNH-containing Tat LTR retrotransposons but lack aRNH, are also denoted on the tree as well as Athila elements and chromoviruses. The names of all aRNH-containing LTR retrotransposons are in bold. The names of aRNH-containing elements identified in the study consist of the host genome acronym, the index number and predicted domain architecture, for example 'Betav-LTR 3097-GAG.PR.RT.gRNH.aRNH.INT'. Abbreviations: aRNH – 'archaeal' type of cellular-like RNH, ENVathila – Envelope-like protein of Athila LTR elements, GAG – Gag protein, gRNH – RNH of Ty3/gypsy LTR retrotransposons, INT – integrase, PR – protease, RT – reverse transcriptase. For species acronyms see Table S1.